

Aggregation and long memory: recent developments*

R. Leipus¹, A. Philippe², D. Puplinskaitė^{1,2} and D. Surgailis¹

June 13, 2013

¹Vilnius University and ²Université de Nantes

Abstract. It is well-known that the aggregated time series might have very different properties from those of the individual series, in particular, long memory. At the present time, aggregation has become one of the main tools for modelling of long memory processes. We review recent work on contemporaneous aggregation of random-coefficient AR(1) and related models, with particular focus on various long memory properties of the aggregated process.

Keywords: Random-coefficient AR(1); contemporaneous aggregation; long memory; infinite variance; mixed moving average; scaling limit; intermediate process; autoregressive random fields; anisotropic long memory; disaggregation.

1 Introduction

In the seminal paper Granger (1980) observed that the covariance function of AR(1) process with random and Beta distributed coefficient may decay very slowly as in ARFIMA process. Indeed, let $X(t) = \sum_{j=0}^{\infty} a^j \zeta(t-j)$ be a stationary solution of AR(1) equation

$$X(t) = aX(t-1) + \zeta(t), \quad t \in \mathbb{Z}, \quad (1.1)$$

where $\{\zeta(t)\} \sim WN(0, \sigma^2)$ is a white noise and $a \in [0, 1)$ is a r.v., independent of $\{\zeta(t)\}$, and having a density $\phi(x)$ regularly varying at the unit root $x = 1$, viz.,

$$\phi(x) \sim c_{\phi}(1-x)^{\beta}, \quad x \nearrow 1, \quad (1.2)$$

where $c_{\phi} > 0$ and $\beta > -1$ some constants. Assume $0 < \beta < 1$ in the rest of this section. Then, as $t \rightarrow \infty$,

$$EX(0)X(t) = \sigma^2 \int_0^1 \frac{x^t}{1-x^2} \phi(x) dx \sim ct^{-\beta}, \quad (1.3)$$

with $c = \sigma^2(c_{\phi}/2)\Gamma(\beta)$ with $\Gamma(\beta) = \int_0^{\infty} e^{-y} y^{\beta-1} dy$. In the case of Beta-type density $\phi(x) = \frac{2}{B(p,q)} x^{2p-1} (1-x^2)^{q-1}$ discussed in Granger (1980), condition (1.2) is satisfied with $\beta = q-1$ and the covariance in (1.3) can be explicitly computed: $EX(0)X(t) = (\sigma^2 \Gamma(q-1)/B(p,q)) \Gamma(p + \frac{t}{2})/\Gamma(p + \frac{t}{2} + q-1)$, leading to the asymptotics in (1.3) since $\Gamma(t)/\Gamma(t+b) \sim t^{-b}$, $t \rightarrow \infty$.

Now suppose that one wants to study a huge and heterogeneous population of dynamic “micro-agents” who all evolve independently of each other according to AR(1) process. Thus, the probability

*The first, third and fourth authors are supported by a grant (No. MIP-13079) from the Research Council of Lithuania.

law of the evolution of a given “micro-agent” is completely determined by the value of the autoregressive parameter a . The heterogeneity of “micro-agents” means that a has a probability density ϕ across the population. The “macroeconomic” variable $\mathfrak{X}_N(t)$ of interest is obtained by averaging the evolutions of N “microeconomic” variables $X_i(t), i = 1, \dots, N$,

$$\mathfrak{X}_N(t) = \frac{1}{A_N} \sum_{i=1}^N X_i(t), \quad t = 0, \pm 1, \dots, \quad (1.4)$$

which are randomly sampled from all “micro-agents” population, where A_N is a normalization and $\{X_i(t)\}, i = 1, 2, \dots$ are i.i.d. copies of (1.1). By the independence of the summands in (1.4) and the classical CLT, it immediately follows (with $A_N = N^{1/2}$) that

$$\mathfrak{X}_N(t) \xrightarrow{\text{fdd}} \mathfrak{X}(t), \quad (1.5)$$

where $\{\mathfrak{X}(t)\}$ is a stationary Gaussian process with the same 2nd moment characteristics as the individual “micro-agents”, i.e.

$$\text{E}\mathfrak{X}(t) = 0, \quad \text{E}\mathfrak{X}(0)\mathfrak{X}(t) = \text{E}X(0)X(t) = \sigma^2 \int_0^1 \frac{x^t}{1-x^2} \phi(x) dx \quad (1.6)$$

and where $\xrightarrow{\text{fdd}}$ denotes the weak convergence of finite-dimensional distributions. In particular, under the assumption (1.2), $0 < \beta < 1$, the covariance $\text{E}\mathfrak{X}(0)\mathfrak{X}(t)$ decays as in (1.3), implying that $\sum_{t \in \mathbb{Z}} |\text{E}\mathfrak{X}(0)\mathfrak{X}(t)| = \infty$, i.e., $\{\mathfrak{X}(t)\}$ has long memory. Notice also that the Gaussian process $\{\mathfrak{X}(t)\}$ in (1.5) is ergodic, in contrast to the non-ergodic random-coefficient individual processes $\{X_i(t)\}$ in (1.4). Since (1.4) refers to summation of (independent) processes at instant time t , the summation procedure (1.4) is called *contemporaneous* or *cross-sectional* aggregation, to be distinguished from *temporal aggregation*, the latter term usually referring to the summation on disjoint successive blocks, or taking partial sums, of a single time series. We see that contemporaneous aggregation of simple dynamic models can result in a long memory process and hence may provide an explanation of the long memory phenomenon observed in many econometric studies (Baillie (1996), Teyssière and Kirman (2007)). The last fact is very important since economic reasons of long memory remain largely unclear and some authors try explain this phenomenon by “spurious long memory” (Lobato and Savin (1998), Mikosch (2003)).

It is also instructive to look at the behavior (1.3) from the “spectral perspective”. The spectral density of the random-coefficient AR(1) process is written as

$$f(y) = \frac{\sigma^2}{2\pi} \int_0^1 \frac{1}{|1 - xe^{iy}|^2} \phi(x) dx = \frac{\sigma^2}{2\pi} \int_0^1 \frac{1}{(1-x)^2 + 4x \sin^2(y/2)} \phi(x) dx, \quad y \in [-\pi, \pi].$$

We see that under (1.2), $0 < \beta < 1$, when $y \searrow 0$, this “aggregated” spectral density behaves like a power function:

$$\begin{aligned} f(y) &\sim \sigma^2 \frac{c_\phi}{2\pi} \int_0^1 \frac{(1-x)^\beta dx}{(1-x)^2 + 4x \sin^2(y/2)} \sim \sigma^2 \frac{c_\phi}{2\pi} \int_0^1 \frac{x^\beta dx}{x^2 + 4 \sin^2(y/2)} \\ &= \sigma^2 \frac{c_\phi}{2\pi (2 \sin(y/2))^{1-\beta}} \int_0^{1/2 \sin(y/2)} \frac{w^\beta dw}{w^2 + 1} \\ &\sim \frac{c_f}{y^{1-\beta}}, \end{aligned} \quad (1.7)$$

where $c_f := \sigma^2 \frac{c_\phi}{2\pi} \int_0^\infty \frac{w^\beta dw}{w^2 + 1}$. The singularity in the low frequency spectrum of $\{X(t)\}$ and $\{\mathfrak{X}(t)\}$ is another indication of long memory of these processes and is not surprising as the power behaviors (1.3) and (1.7) are known to be roughly equivalent. Relation (1.7) seems to be well-known to physicists and

even experimentally observed, relating aggregation and long memory to the vast area in physics called $1/f$ noise (see, e.g., the article by Ward and Greenwood (2007) in Scholarpedia). Another manifestation of long memory for the aggregated process in (1.7) is the convergence of normalized partial sums of $\{\mathfrak{X}(t)\}$ to a fractional Brownian motion (Zaffaroni (2004)). See Sec. 2 for the related notion of distributional long memory.

Following Granger (1980), various aspects of aggregation were discussed in the literature, see Gonçalves and Gouriéroux (1988), Zaffaroni (2004, 2007a), Oppenheim and Viano (2004), Leipus *et al.* (2006), Celov *et al.* (2007, 2010), Jirak (2011). Oppenheim and Viano (2004) considered aggregation of general AR($2p$) processes with $2p$ random coefficients, under the condition that the characteristic polynomial has the form

$$A(z) = (1 - \alpha_1 z)(1 + \alpha_2 z) \prod_{k=3}^{p+1} (1 - \alpha_k e^{i\theta_k} z)(1 - \alpha_k e^{-i\theta_k} z),$$

where $\alpha_1, \dots, \alpha_{p+1}$ are i.i.d. random variables with probability densities as in (1.2), with (possibly) different β 's, and $\theta_3, \dots, \theta_{p+1}$ are fixed. They described the singularities of the spectral density and the asymptotics of the covariance function of the corresponding aggregated Gaussian process, which may contain a seasonal component. Zaffaroni (2004) considered a general aggregation scheme of random-coefficient AR(1) with idiosyncratic and common components. Several papers (Ding and Granger (1996), Zaffaroni (2007a), Zaffaroni (2007b), Leipus and Viano (2002), Kazakevičius *et al.* (2004), Giraitis *et al.* (2010)) extended the aggregation scheme to ARCH-type heteroskedastic processes with random coefficients and common innovations, with a particular emphasis on long memory behavior. The disaggregation problem of reconstructing the mixing distribution from the spectral density or observed sample from the aggregated process and its statistical aspects were studied Dacunha-Castelle and Fermin (2006), Celov *et al.* (2007), Leipus *et al.* (2006), Celov *et al.* (2010), Jirak (2011), Robinson (1978), Beran *et al.* (2010) and elsewhere. Some aspects of aggregation and disaggregation are also discussed in the recent monograph Beran and Kulik (2013).

The aim of the present paper is to review recent developments on aggregation and long memory. Sec. 2 discusses the case of AR(1) processes with infinite variance. Sec. 3 extends the aggregation procedure to the triangular array model. Sec. 4 studies the joint temporal and contemporaneous aggregation of AR(1) processes. Sec. 5 considers the aggregation procedure for random fields. Sec. 6 reviews some results related to statistical inference for the mixing density ϕ .

2 Aggregation of AR(1) processes with infinite variance

The approach of Granger (1980) presented Sec. 1 can be extended to the case of infinite variance. Let

$$X(t) = \sum_{j=0}^{\infty} a^j \zeta(t-j) \tag{2.1}$$

be a stationary solution of the AR(1) equation (1.1) with random coefficient $a \in [0, 1)$ and i.i.d. innovations $\{\zeta(t)\}$ with *infinite variance* $\text{Var}(\zeta(t)) = \infty$, independent of r.v. a , and belonging to the domain of attraction of α -stable law, $0 < \alpha < 2$, in the sense that

$$\frac{1}{N^{1/\alpha}} \sum_{i=1}^N \zeta(i) \xrightarrow{d} Z, \tag{2.2}$$

where Z is an α -stable r.v. Let $\{X_i(t)\}, i = 1, 2, \dots$ be independent copies of (2.1), and define the aggregated process as

$$\mathfrak{X}_N(t) := \frac{1}{N^{1/\alpha}} \sum_{i=1}^N X_i(t), \quad t \in \mathbb{Z}. \quad (2.3)$$

The problem is to determine the limit process $\{\mathfrak{X}(t)\}$ of (2.3), in the sense of (1.5), and then we describe its properties, in particular, long memory properties.

The above questions were studied in Puplinskaitė and Surgailis (2010). Note that the AR(1) series in (2.1) converges conditionally a.s. for every $a \in [0, 1)$, and also unconditionally a.s. if the distribution of a satisfies the additional condition $E(1-a)^{-1} < \infty$. For regularly varying ϕ as in (1.2), the last condition is equivalent to $\beta > 0$. It is shown in Puplinskaitė and Surgailis (2010, Thm. 2.1) that in the case $\beta > 0$ the limit of $\{\mathfrak{X}_N(t)\}$ exists and is written as a stochastic integral

$$\mathfrak{X}(t) := \sum_{s \leq t} \int_{[0,1)} x^{t-s} M_{\alpha,s}(dx), \quad (2.4)$$

where $\{M_{\alpha,s}, s \in \mathbb{Z}\}$ are i.i.d. copies of an α -stable random measure on $[0, 1)$ with distribution Z in (2.2) and control measure equal to the mixing distribution Φ (i.e., the distribution of the r.v. a). Recall that a family $M = \{M(A), A \in \mathcal{B}_0(S)\}$ of r.v.'s indexed by sets $A \in \mathcal{B}_0(S) := \{A \in \mathcal{B}(S) : \mu(A) < \infty\}$ of a measure space $(S, \mathcal{B}(S), \mu)$ with a σ -finite measure μ is called a *random measure with distribution W and control measure μ* (where W is an infinitely divisible r.v.) if for any disjoint sets $A_i \in \mathcal{B}_0(S)$, $i = 1, \dots, n$, $n \geq 1$ r.v.'s $M(A_i), i = 1, \dots, n$ are independent, $M(\cup_{i=1}^n A_i) = \sum_{i=1}^n M(A_i)$, and

$$Ee^{i\theta M(A)} = (Ee^{i\theta W})^{\mu(A)}, \quad \forall A \in \mathcal{B}_0(S), \quad \forall \theta \in \mathbb{R}.$$

Integration with respect to stable and general infinitely divisible random measures is discussed in Samorodnitsky and Taqqu (1994), Rajput and Rosinski (1989) and other texts. The process in (2.4) is a particularly simple case of the so-called *mixed stable moving averages* introduced in Surgailis *et al.* (1993). It is stationary, ergodic, and has α -stable finite-dimensional distributions. The representation (2.4) of the limit aggregated process holds also in the finite-variance case $\alpha = 2$, yielding a stationary Gaussian process with zero mean and covariance

$$\text{Cov}(\mathfrak{X}(0), \mathfrak{X}(t)) = \sigma^2 \sum_{s \leq 0} \int_{[0,1)} x^{-s} x^{t-s} \Phi(dx) = \sigma^2 E \left[\frac{a^t}{1-a^2} \right], \quad t = 0, 1, \dots,$$

cf. (1.3). The mixed stable moving average in (2.4) can be regarded as a limiting “superposition” $\sum_{a_i \in [0,1)} \mathcal{X}(t; a_i)$ of independent α -stable AR(1) processes $\mathcal{X}(t; a_i) = \sum_{s \leq t} a_i^{t-s} \mathcal{M}_{\alpha,i}(t-s)$ with α -stable innovations $\{\mathcal{M}_{\alpha,i}(s)\}$. Although each $\{\mathcal{X}(t; a_i)\}$ is geometrically mixing and hence short memory, the dependence in $\{\mathcal{X}(t; a_i)\}$ increases when $a_i \nearrow 1$ approaches the “unit root” $a = 1$. It turns out that the limiting “superposition” of these processes and the mixed moving average in (2.4) may have long (or short) memory, depending on the concentration of the a_i 's near $a = 1$, or the parameter β in (1.2).

Before addressing the question about long memory of the infinite variance process in (2.4), let us note that the above mentioned convergence of (2.3) to (2.4) does not hold in the case of negative exponent $-1 < \beta < 0$. It turns out that in the latter case, the limit aggregated process does not depend on t and is an $\alpha(1 + \beta)$ -stable r.v. (random constant):

$$\frac{1}{N^{1/(\alpha(1+\beta))}} \sum_{i=1}^N X_i(t) \xrightarrow{\text{fdd}} \tilde{Z}, \quad (2.5)$$

see Puplinskaitė and Surgailis (2010, Prop. 2.3). In the case $\alpha = 2$, a similar result was noted in Zaffaroni (2004). Note that the normalizing exponents $1/\alpha$ in (2.3) and $1/(\alpha(1 + \beta))$ in (2.5) are different and

$1/(\alpha(1+\beta)) \rightarrow \infty$ as $\beta \rightarrow -1$. Therefore, $\beta = 0$ is a critical point resulting in completely different limits of the aggregated process in the cases $\beta > 0$ and $\beta < 0$. The fact that the limit is degenerate in the latter case can be intuitively explained as follows. It is clear that, with β decreasing, the dependence increases in the random-coefficient AR(1) process $\{X(t)\}$, as well as in the limiting aggregated process $\{\mathfrak{X}(t)\}$. For negative $\beta < 0$, the dependence in the aggregated process becomes extremely strong so that the limit process is degenerate and completely dependent.

Long memory properties of the limit aggregated process. Clearly, the usual definitions of long memory in terms of covariance/spectrum do not apply in infinite variance case. Alternative notions of long memory which are applicable to infinite-variance processes have been proposed in the literature. Astrauskas (1983) was probably the first to rigorously study long memory for such processes in terms of the rate at which the bivariate characteristic function at distant lags factorizes into the product of two univariate characteristic functions. Related characteristics such as codifference are discussed in Samorodnitsky and Taqqu (1994). Some characteristics of dependence (covariation, α -covariance) for stable processes expressed in terms of the spectral measure were studied in Samorodnitsky and Taqqu (1994) and Paulauskas (2013). Heyde and Yang (1997) defined the long-range dependence (sample Allen variance) (LRD(SAV)) property in terms of the limit behavior of squared studentized sample mean.

Probably, the most useful and universal definition of long/short memory was given by Cox (1984). Assume that $\{Y(t), t \in \mathbb{Z}\}$ is a strictly stationary process series and there exist some constants $D_n \rightarrow \infty$ ($n \rightarrow \infty$) and B_n and a nontrivial stochastic process $J = \{J(\tau), \tau \geq 0\} \not\equiv 0$ such that

$$\frac{1}{D_n} \sum_{s=1}^{[n\tau]} (Y(s) - B_n) \xrightarrow{\text{fdd}} J(\tau). \quad (2.6)$$

According to Cox (1984), if the limit process J has *independent increments*, the series $\{Y(t)\}$ is said to have *distributional short memory*, while in the converse case when J has *dependent increments*, the series $\{Y(t)\}$ is said to have *distributional long memory*.

The above definition has several advantages. First of all, it does not depend on any moment assumptions since finite and infinite variance processes are treated from the same angle. Secondly, according to the classical Lamperti's theorem (see Lamperti (1962)), in the case of (2.6) there exists a number $H > 0$ such that the normalizing constants A_n grow as n^H (modulus a slowly varying factor), while the limit random process J is H -self-similar and has stationary increments (H -sssi). A faster growth of normalization A_n means stronger dependence and therefore H is a quantitative indicator of the degree of dependence in $\{Y(t)\}$. The characterization of short memory through (2.6) is very robust and essentially reduces to Lévy stable behavior since all sssi processes with independent increments are stable Lévy processes. We emphasize that the above definition of long/short memory requires identification of the partial sums limit which is sometimes not easy. On the other hand, partial sums play a very important role in statistical inference, especially under long memory, see e.g. Giraitis *et al.* (2012), and hence finding partial sums limit is very natural for understanding the dependence structure of a given process and subsequent applications of statistical nature.

With the above discussion in mind, the question arises what is the partial sums limit of the aggregated process in (2.4)? This question is answered in Puplinskaitė and Surgailis (2010, Thm. 3.1) saying that for $1 < \alpha \leq 2$ and $0 < \beta < \alpha - 1$,

$$\frac{1}{n^{1-(\beta/\alpha)}} \sum_{t=1}^{[n\tau]} \mathfrak{X}(t) \xrightarrow{\text{fdd}} \Lambda_{\alpha,\beta}(\tau) := \int_{\mathbb{R}_+ \times \mathbb{R}} (\mathfrak{f}(x, \tau - s) - \mathfrak{f}(x, -s)) Z_\alpha(dx, ds), \quad (2.7)$$

where

$$f(x, t) := \begin{cases} (1 - e^{-xt})/x, & \text{if } x > 0 \text{ and } t > 0, \\ 0, & \text{otherwise,} \end{cases} \quad (2.8)$$

and $Z_\alpha(dx, ds)$ is an α -stable random measure on $\mathbb{R}_+ \times \mathbb{R}$ with control measure $c_\phi x^\beta dx ds$ and distribution Z , where Z is defined at (2.2). The random process $\Lambda_{\alpha, \beta}$ in (2.7) is well-defined for $1 < \alpha \leq 2$, $0 < \beta < \alpha - 1$ and is H -sssi with self-similarity index $H = 1 - \frac{\beta}{\alpha} \in (\frac{1}{\alpha}, 1)$. Moreover, $\Lambda_{\alpha, \beta}$ has a.s. continuous paths, α -stable finite dimensional distributions and stationary and dependent increments. In particular, $\Lambda_{2, \beta}$ is a fractional Brownian motion with $H = 1 - \frac{\beta}{2} \in (\frac{1}{2}, 1)$. The process $\Lambda_{\alpha, \beta}$ is also different from linear fractional Lévy motion (see, e.g., Samorodnitsky and Taqqu (1994) for a detailed discussion of the latter process).

Similarly as $\{\mathfrak{X}(t)\}$ in (2.4) can be regarded as a “continuous superposition” of AR(1) processes, the limit process $\Lambda_{\alpha, \beta}$ in (2.7) can be regarded as a “continuous superposition” of (integrated) Ornstein-Uhlenbeck processes $z_\alpha(\tau; x)$ defined as

$$z_\alpha(\tau; x) := \int_0^\tau du \int_{\mathbb{R}} e^{-x(u-s)} \mathbf{1}(u > s) dZ_\alpha(s) = \int_{\mathbb{R}} (f(x, \tau-s) - f(x, -s)) dZ_\alpha(s), \quad \tau \geq 0, x > 0, \quad (2.9)$$

where $\{Z_\alpha(s), s \geq 0\}$ is an α -stable Lévy process with independent increments. Note that for each $x > 0$, the process $\{z_\alpha(\tau; x)\}$ is a.s. continuously differentiable on \mathbb{R} and its derivative $z'_\alpha(\tau; x) = dz_\alpha(\tau; x)/d\tau$ satisfies the Langevin equation

$$dz'_\alpha(\tau; x) = -xz'_\alpha(\tau; x)d\tau + dZ_\alpha(\tau).$$

In the case $\alpha = 2$, $Z_2 = B$ is a usual Brownian motion and $z_2(\tau; x)$ is a Gaussian Ornstein-Uhlenbeck process. The corresponding representation of $\Lambda_{2, \beta}$ in (2.7) may be termed the *Ornstein-Uhlenbeck representation of fractional Brownian motion*, and the process $\Lambda_{\alpha, \beta}$, $1 < \alpha < 2$ its stable counterpart. A related class of stationary infinitely divisible processes with long memory is discussed in Barndorff-Nielsen (2001).

The condition $1 < \alpha \leq 2$, $0 < \beta < \alpha - 1$ for the convergence in (2.7) is sharp and cannot be weakened. In particular, for $\beta > \alpha - 1$ the partial sums process $n^{-1/\alpha} \sum_{t=1}^{[n\tau]} \mathfrak{X}(t)$ tends to an α -stable Lévy process with independent increments (see Puplinskaitė and Surgailis (2010, Thm. 3.1)). Therefore, large sample behaviors of $\{\mathfrak{X}(t)\}$ for $0 < \beta < \alpha - 1$ and $\beta > \alpha - 1$ are markedly different. Following the above terminology, the aggregated AR(1) process $\{\mathfrak{X}(t)\}$ has distributional long memory if $0 < \beta < \alpha - 1$ and distributional short memory if $\beta > \alpha - 1$. In the last case, assumption (1.2) on the mixing distribution can be substantially relaxed.

Puplinskaitė and Surgailis (2010) also studied other characterizations of long memory of the aggregated AR(1) process $\{\mathfrak{X}(t)\}$ with infinite variance (the LRD(SAV) property, the decay rate of codiference). A curious characterization of long memory in terms of the asymptotic behaviour of the *ruin probability* in a discrete time risk insurance model with α -stable claims $\{\mathfrak{X}(t)\}$ in (2.4) is obtained in Perilioglu and Puplinskaitė (2013), following the characterization studied in Mikosch and Samorodnitsky (2000). All these results agree with the above characterization in terms of the partial sums process, in the sense that $\beta = \alpha - 1$ is the boundary between long memory and short memory in $\{\mathfrak{X}(t)\}$.

Finally, let us note that the general aggregation scheme of random-coefficient autoregressive processes discussed in Zaffaroni (2004) includes the case of common component, or *common innovations*. Aggregation of infinite variance AR(1) processes with common innovations was studied in Puplinskaitė and Surgailis (2009). In this case, the limit aggregated process, say $\{\tilde{\mathfrak{X}}(t)\}$, exists under normalization $1/N$ and is a moving average with the same innovations as the original AR(1) series (2.1) and the moving average coefficients given by the expectations $Ea^j = \int_{[0,1)} x^j \Phi(dx)$. By a similar argument as in (1.3), we

have that $Ea^j \sim \text{const } j^{-\beta-1}$ for $\beta > -1$. Therefore for $1/\alpha < \beta < 0$, $\{\tilde{\mathfrak{X}}(t)\}$ is a well-defined long memory moving average with infinite variance and nonsummable coefficients $\sum_{j=0}^{\infty} Ea^j = \infty$. Puplinskaitė and Surgailis (2009) investigated various long memory properties of $\{\tilde{\mathfrak{X}}(t)\}$, including the convergence of its partial sums to a linear fractional Lévy motion.

3 Aggregation of AR(1) processes: triangular array innovations

The contemporaneous aggregation scheme of Sec. 2 can be generalized by assuming that the innovations depend on N , constituting a triangular array of i.i.d. r.v.'s. Such aggregation scheme was studied in Philippe *et al.* (2013). Let $\{X_i^{(N)}(t)\}$, $i = 1, \dots, N$ be i.i.d. copies of random-coefficient AR(1) process

$$X^{(N)}(t) = aX^{(N)}(t-1) + \zeta^{(N)}(t), \quad t \in \mathbb{Z}, \quad (3.1)$$

where $\{\zeta^{(N)}(t), t \in \mathbb{Z}\}$, $N = 1, 2, \dots$ is a triangular array of i.i.d. random variables in the domain of attraction of an infinitely divisible law W :

$$\sum_{i=1}^N \zeta^{(N)}(i) \xrightarrow{d} W, \quad (3.2)$$

and $a \in [0, 1)$ is a r.v., independent of $\{\zeta^{(N)}(t), t \in \mathbb{Z}\}$. The limit aggregated process $\{\mathfrak{X}(t), t \in \mathbb{Z}\}$ is defined as the limit in distribution:

$$\sum_{i=1}^N X_i^{(N)}(t) \xrightarrow{\text{fdd}} \mathfrak{X}(t). \quad (3.3)$$

Sec. 2 corresponds to the particular case of (3.1)–(3.2), viz., $\zeta^{(N)}(t) = N^{-1/\alpha} \zeta(t)$, where $\{\zeta(t), t \in \mathbb{Z}\}$ are i.i.d. r.v.'s in the domain of (normal) attraction of α -stable law W , $0 < \alpha \leq 2$. In particular, for $\alpha = 2$ or $W \sim \mathcal{N}(0, \sigma^2)$ the last condition is equivalent to $E\zeta(t) = 0$ and $\sigma^2 = EW^2 < \infty$.

One of the main results of Philippe *et al.* (2013) says that under mild additional conditions the limit in (3.3) exists and is written as a mixed infinitely divisible (ID) moving average (see the terminology in Rajput and Rosinski (1989)):

$$\mathfrak{X}(t) = \sum_{s \leq t} \int_{[0,1)} x^{t-s} M_{W,s}(dx), \quad t \in \mathbb{Z}, \quad (3.4)$$

where $\{M_{W,s}, s \in \mathbb{Z}\}$ are i.i.d. copies of an ID random measure M_W on $[0, 1)$ with control measure $\Phi(dx) = P(a \in dx)$ and the distribution W in (3.2). Recall that the last distribution is uniquely determined by its Lévy characteristics (μ, σ, π) (the characteristic triplet) since

$$V(\theta) := \log Ee^{i\theta W} = \int_{\mathbb{R}} (e^{i\theta x} - 1 - i\theta x \mathbf{1}(|x| \leq 1)) \pi(dx) - \frac{1}{2} \theta^2 \sigma^2 + i\theta \mu, \quad (3.5)$$

where $\mu \in \mathbb{R}$, $\sigma \geq 0$ and π is a Lévy measure, see Samorodnitsky and Taqqu (1994), Sato (1999).

Philippe *et al.* (2013) discuss distributional long/short memory properties of the aggregated process $\{\mathfrak{X}(t)\}$ in (3.4) with finite variance and a mixing density as in (1.2). The finite variance assumption is equivalent to $E(1-a)^{-1} < \infty$ and

$$\sigma_W^2 := \text{Var}(W) < \infty \iff \int_{\mathbb{R}} x^2 \pi(dx) < \infty. \quad (3.6)$$

Note that (3.6) excludes the α -stable case discussed in the previous sec. Under (3.6) the covariance function of (3.4) is written as in the Gaussian case (1.6), with σ^2 replaced by σ_W^2 . Therefore the covariance

asymptotics in (1.3) applies also for the process $\{\mathfrak{X}(t)\}$ in (3.4), yielding

$$\text{Var}\left(\sum_{t=1}^n \mathfrak{X}(t)\right) \sim Cn^{2-\beta}, \quad 0 < \beta < 1. \quad (3.7)$$

From (3.7) and the linear structure of $\{\mathfrak{X}(t)\}$ one might expect a Gaussian (fractional Brownian motion) limit behavior of the partial sums process $S_n(\tau) = \sum_{t=1}^{\lfloor n\tau \rfloor} \mathfrak{X}(t)$.

However, as it turns out, the Gaussian scenario for $\{S_n(\tau)\}$ is valid only if $\sigma > 0$ in (3.5), or the Gaussian component is present in the ID r.v. W . Else (i.e., when $\sigma = 0$), the behavior of the Lévy measure π at the origin plays a dominant role. Assume that there exist $\alpha_0 > 0$ and $c_0^\pm \geq 0, c_0^+ + c_0^- > 0$ such that

$$\lim_{x \searrow 0} x^{\alpha_0} \pi(x, \infty) = c_0^+, \quad \lim_{x \searrow 0} x^{\alpha_0} \pi(-\infty, -x) = c_0^-. \quad (3.8)$$

It is proved in Philippe *et al.* (2013) that under conditions (1.2), (3.6), and (3.8), partial sums $\{S_n(\tau)\}$ of $\{\mathfrak{X}(t)\}$ in (3.4) may exhibit at least four different limit behaviors, depending on parameters β, σ , and α_0 . The four parameter regions and the limit behaviors in the f.d.d. sense are described in (i)–(iv) below.

- (i) $0 < \beta < 1, \sigma > 0$. In this region, $n^{(\beta/2)-1}S_n(\tau)$ tends to a fractional Brownian motion with Hurst parameter $H = 1 - (\beta/2)$.
- (ii) $0 < \beta < 1, \sigma = 0, 1 + \beta < \alpha_0 < 2$. In this region, $n^{(\beta/\alpha_0)-1}S_n(\tau)$ tends to the α_0 –stable self-similar process $\Lambda_{\alpha_0, \beta}$ defined in (2.7).
- (iii) $0 < \beta < 1, \sigma = 0, 0 < \alpha_0 < 1 + \beta$. In this region, $n^{-1/(1+\beta)}S_n(\tau)$ tends to a $(1 + \beta)$ –stable Lévy process with independent increments.
- (iv) $\beta > 1$. In this region, $n^{-1/2}S_n(\tau)$ tends to a Brownian motion.

See Philippe *et al.* (2013) for precise formulations. Accordingly, the process $\{\mathfrak{X}(t)\}$ in (3.4) has distributional long memory in regions (i) and (ii) and distributional short memory in regions (iii) and (iv). As α_0 increases from 0 to 2, the Lévy measure in (3.8) increases its “mass” near the origin, the limiting case $\alpha_0 = 2$ corresponding to $\sigma > 0$ or a positive “mass” at 0. We see from (i)–(ii) that distributional long memory is related to α_0 being large enough, or small jumps of the random measure M_W having sufficient high intensity. Note that the critical exponent $\alpha_0 = 1 + \beta$ separating the long and short memory “regimes” in (ii) and (iii) decreases with β , which is quite natural since smaller β means the mixing distribution putting more weight near the unit root $a = 1$.

Let us note that an α –stable limit behavior of partial sums of stationary finite variance processes is not unusual under long memory. See, e.g., Willinger *et al.* (1997), Mikosch *et al.* (2002), Leipus and Surgailis (2003) and the references herein. On the other hand, these papers focus on heavy-tailed duration models in which case the limit α –stable process has *independent increments* as a rule. The situation when an infinite variance limit process with *dependent increments* arises from partial sums of a finite variance process as in (ii) above seems rather new.

4 Joint temporal and contemporaneous aggregation of AR(1) processes

The aggregation procedures discussed in the previous sec. extend in a natural way to the (large scale) joint temporal and contemporaneous aggregation. In the latter frame, we are interested in the limit

behavior of the double sums

$$S_{N,n}(\tau) := \sum_{i=1}^N \sum_{t=1}^{[n\tau]} X_i(t), \quad \tau \geq 0, \quad (4.1)$$

where $\{X_i(t)\}, i = 1, \dots, N$ are the same random-coefficient AR(1) processes as in (1.4). The sum in (4.1) represents joint temporal and contemporaneous aggregate of N individual AR(1) evolutions at time scale n . The main question is the joint aggregation limit $\lim_{N,n \rightarrow \infty} A_{N,n}^{-1} S_{N,n}(\tau)$, in distribution, where $A_{N,n}$ are some normalizing constants and both N and n increase to infinity, possibly at different rate. This question was studied in Pilipauskaitė and Surgailis (2013). The last paper also discussed the iterated limits of $A_{N,n}^{-1} S_{N,n}(\tau)$ when first $n \rightarrow \infty$ and then $N \rightarrow \infty$, or vice-versa. Remark that the discussion in the previous sec. refers to the latter iterated limit as $N \rightarrow \infty$ first, followed by $n \rightarrow \infty$. Similar questions for some network traffic models were studied in Willinger *et al.* (1997), Mikosch *et al.* (2002), Gaigalas and Kaj (2003), Pipiras *et al.* (2004), Dombry and Kaj (2011) and other papers. In these papers, the role of AR(1) processes $\{X_i(t)\}$ in (4.1) play independent ON/OFF processes or M/G/ ∞ queues with heavy-tailed activity periods.

Let us describe the main results in Pilipauskaitė and Surgailis (2013). They refer to the random-coefficient AR(1) process with i.i.d. innovations having zero mean and variance $\sigma^2 < \infty$, and a mixing density as (1.2). Let N, n increase simultaneously so as

$$\frac{N^{1/(1+\beta)}}{n} \rightarrow \mu \in [0, \infty], \quad (4.2)$$

leading to the following three cases:

$$\text{Case (j)} : \mu = \infty, \quad \text{Case (jj)} : \mu = 0, \quad \text{Case (jjj)} : 0 < \mu < \infty. \quad (4.3)$$

Following the terminology in Mikosch *et al.* (2002) and Pipiras *et al.* (2004), we call Cases (j), (jj), and (jjj) the “fast growth condition”, the “slow growth condition”, and the “intermediate growth condition”, respectively, since they reflect how fast N grows with n . The main result of Pilipauskaitė and Surgailis (2013) says that under (4.2), the “simultaneous limit” of $S_{N,n}(\tau)$ exist and are different in all three Cases (j)–(jjj).

Case (j) (the “fast growth condition”): For any $0 < \beta < 1$,

$$N^{-1/2} n^{-1+(\beta/2)} S_{N,n}(\tau) \xrightarrow{\text{fdd}} B_{1-\beta/2}(\tau), \quad (4.4)$$

where $\{B_{1-\beta/2}(\tau)\}$ is a fractional Brownian motion with $H = 1 - \beta/2 \in (1/2, 1)$.

Case (jj) (the “slow growth condition”): For any $-1 < \beta < 1$,

$$N^{-1/(1+\beta)} n^{-1/2} S_{N,n}(\tau) \xrightarrow{\text{fdd}} \mathcal{W}_\beta(\tau), \quad (4.5)$$

where $\{\mathcal{W}_\beta(\tau)\}$ is a sub-Gaussian $(1 + \beta)$ -stable process defined as $\mathcal{W}_\beta(\tau) = W_\beta^{1/2} B(\tau)$, $\tau \geq 0$, where $W_\beta > 0$ is a $(1 + \beta)/2$ -stable totally skewed r.v. and $\{B(\tau), \tau \geq 0\}$ is a standard Brownian motion independent of W_β (see, e.g., Samorodnitsky and Taqqu (1994)).

Case (jjj) (the “intermediate growth condition”): For any $-1 < \beta < 1$

$$N^{-1/(1+\beta)} n^{-1/2} S_{N,n}(\tau) \xrightarrow{\text{fdd}} \mu^{1/2} \mathcal{Z}_\beta(\tau/\mu), \quad (4.6)$$

where the limit process $\{\mathcal{Z}_\beta(\tau), \tau \geq 0\}$ is defined through finite-dimensional characteristic function:

$$\mathbb{E} \exp \left\{ i \sum_{j=1}^m \theta_j \mathcal{Z}_\beta(\tau_j) \right\} = \exp \left\{ c_\phi \int_{\mathbb{R}_+} \left(e^{-(\sigma^2/2) \int_{\mathbb{R}} \left(\sum_{j=1}^m \theta_j (\mathfrak{f}(x, \tau_j - s) - \mathfrak{f}(x, -s)) \right)^2 ds} - 1 \right) x^\beta dx \right\}, \quad (4.7)$$

where $\theta_j \in \mathbb{R}, \tau_j \in \mathbb{R}_+, j = 1, \dots, m, m \geq 1$, and \mathbf{f} is given in (2.8).

Let us give some comments on the above results. The convergence in Case (j) (4.4) is very natural in view of (2.7) (with $\alpha = 2$ and $\Lambda_{2,\beta} = B_{1-\beta/2}$ a fractional Brownian motion). Indeed, under the “fast growth condition” we expect that $N^{-1/2}n^{-1+(\beta/2)}S_{N,n}(\tau)$ can be approximated by $n^{-1+(\beta/2)}\sum_{t=1}^{[n\tau]}\mathfrak{X}(t)$, which converges to $B_{1-\beta/2}(\tau)$ as it happens in the case when $N \rightarrow \infty$ followed by $n \rightarrow \infty$. The limit in Case (jj) (4.5) can be also easily explained since “individual” partial sums $\sum_{t=1}^{[n\tau]}X_i(t)$ behave as Brownian motions with random variances $(1-a_i)^{-2}$ having infinite expectation and a heavy tailed distribution with tail parameter $(1+\beta)/2 \in (0, 2)$: $P((1-a_i)^{-2} > x) = P(a_i > 1 - 1/\sqrt{x}) \sim Cx^{-(\beta+1)/2}, x \rightarrow \infty$. The sum of such independent “random-variance” Brownian motions behaves as a sub-Gaussian process \mathcal{W}_β in (4.5). Particularly interesting is the limit process \mathcal{Z}_β arising under “intermediate scaling” in Case (jjj). It is shown in Pilipauskaitė and Surgailis (2013) that \mathcal{Z}_β admits a stochastic integral with respect to a Poisson random measure on the product space $\mathbb{R} \times C(\mathbb{R})$ with mean $\psi_1 x^\beta dx \times P_B$, where P_B is the Wiener measure on $C(\mathbb{R})$ and enjoys several “intermediate” properties between the limits in (j) and (jj). According to (4.7), \mathcal{Z}_β has infinitely divisible finite-dimensional distributions and stationary increments, but is neither self-similar nor stable. For $0 < \beta < 1$ the process \mathcal{Z}_β has finite variance and the covariance equal to that of a fractional Brownian motion. These results can be compared to Mikosch *et al.* (2002), Gaigalas and Kaj (2003), Pipiras *et al.* (2004), Dombry and Kaj (2011). In particular, Mikosch *et al.* (2002) discuss the “total accumulated input” $\mathcal{A}_{n,N}(\tau) := \int_0^{n\tau} \sum_{i=1}^N W_i(t)dt$ from N independent “sources” $\{W_i(t), i = 1, \dots, N; t \geq 0\}$ at time scale n . The aggregated inputs are i.i.d. copies of ON/OFF process $\{W_t, t \geq 0\}$, alternating between 1 and 0 and taking value 1 if t is in an ON-period and 0 if t is in an OFF-period, the ON- and OFF-periods forming a stationary renewal process having heavy-tailed lengths with respective tail parameters $\alpha_{\text{on}}, \alpha_{\text{off}} \in (1, 2), \alpha_{\text{on}} < \alpha_{\text{off}}$. The role of condition (4.2) is played in the above paper by

$$\frac{n}{N^{\alpha_{\text{on}}-1}} \rightarrow \mu \in [0, \infty], \quad (4.8)$$

leading to the three cases analogous to (4.3):

$$\text{Case (j')} : \mu = 0, \quad \text{Case (jj')} : \mu = \infty, \quad \text{Case (jjj')} : 0 < \mu < \infty. \quad (4.9)$$

The limit of (normalized) “input” $\mathcal{A}_{n,N}(\tau)$ in Case (j') (the “slow growth condition”) and Case (jj') (the “fast growth condition”) was obtained in Mikosch *et al.* (2002), as an α -stable Lévy process and a fractional Brownian motion, respectively. The “intermediate” limit in Case (jjj') was identified in Gaigalas and Kaj (2003), Gaigalas (2006), Dombry and Kaj (2011) who showed that this process can be regarded as a “bridge” between the limiting processes in Cases (j') and (jj'), and can be represented as a stochastic integral with respect to a Poisson random measure on $\mathbb{R}_+ \times \mathbb{R}$. A common feature to the above research and Pilipauskaitė and Surgailis (2013) is the fact the partial sums of the individual processes with finite variance tend to an infinite variance process, thus exhibiting an increase of variability. See also Leipus and Surgailis (2003), Leipus *et al.* (2005). These analogies raise interesting open questions about extension of the joint temporal-contemporaneous aggregation scheme to general independent processes with covariance long memory and stable behavior of partial sums.

5 Aggregation of autoregressive random fields

The idea of aggregation naturally extends to spatial autoregressive models (Lavancier (2007), Lavancier (2011), Puplinskaitė and Surgailis (2013), Lavancier *et al.* (2013)). Following Puplinskaitė and Surgailis (2013), consider a nearest-neighbor autoregressive random field $\{X(t, s)\}$ on \mathbb{Z}^2 satisfying the difference

equation

$$X(t, s) = \sum_{|u|=|v|=1} a(u, v)X(t+u, s+v) + \zeta(t, s), \quad (t, s) \in \mathbb{Z}^2, \quad (5.1)$$

where $\{\zeta(t, s), (t, s) \in \mathbb{Z}^2\}$ are i.i.d. r.v.'s whose generic distribution $\zeta \in D(\alpha)$ belongs to the domain of (normal) attraction of α -stable law, $1 < \alpha \leq 2$, and $a(t, s) \geq 0$, $|t| = |s| = 1$ are *random* coefficients, independent of $\{\zeta(t, s)\}$ and satisfying the condition $A := \sum_{|t|=|s|=1} a(t, s) < 1$ a.s. for the existence of a stationary solution of (5.1). The stationary solution of (5.1) is given by the convergent series

$$X(t, s) = \sum_{(u, v) \in \mathbb{Z}^2} g(t-u, s-v, a)\zeta(u, v), \quad (t, s) \in \mathbb{Z}^2, \quad (5.2)$$

where $g(t, s, a)$, $(t, s) \in \mathbb{Z}^2$, $a = (a(t, s), |t| = |s| = 1)$, is the (random) lattice Green function solving the equation $g(t, s, a) - \sum_{|u|=|v|=1} a(u, v)g(t+u, s+v, a) = \delta(t, s)$, where $\delta(t, s)$ is the delta function. Under the condition $E(1-A)^{-1} < \infty$, the series in (5.2) converges unconditionally in L^p for any $1 < p < \alpha$ (Lavancier *et al.* (2013)). In the finite variance case $\alpha = 2$, the stationary solution (5.2) can be defined via spectral representation:

$$X(t, s) = \int_{[-\pi, \pi]^2} e^{i(tx+sy)} \hat{g}(x, y, a) Z(dx, dy), \quad (5.3)$$

where $\hat{g}(x, y, a) = (1 - \sum_{|t|=|s|=1} a(t, s)e^{i(tx+sy)})^{-1}$ is the Fourier transform of $g(t, s, a)$ and $Z(dx, dy)$ is the random measure satisfying $\int_{[-\pi, \pi]^2} e^{i(tx+sy)} Z(dx, dy) = \zeta(t, s)$. The spectral density of (5.3) is written similarly to one-dimensional case

$$f(x, y) = \frac{\sigma^2}{(2\pi)^2} E|\hat{g}(x, y, a)|^2, \quad (x, y) \in [-\pi, \pi]^2, \quad \sigma^2 = E\zeta^2. \quad (5.4)$$

Let $\{X_i(t, s)\}$, $i = 1, 2, \dots$ be independent copies of (5.2). The aggregated field $\{\mathfrak{X}(t, s)\}$ is defined as the limit in distribution:

$$N^{-1/\alpha} \sum_{i=1}^N X_i(t, s) \xrightarrow{\text{fdd}} \mathfrak{X}(t, s), \quad (t, s) \in \mathbb{Z}^2. \quad (5.5)$$

Under mild additional conditions, Puplinskaitė and Surgailis (2013) prove that the limit in (5.5) exists and is written as a stochastic integral

$$\mathfrak{X}(t, s) = \sum_{(u, v) \in \mathbb{Z}^2} \int_{\mathbf{A}} g(t-u, s-v, a) M_{u, v}(da), \quad (t, s) \in \mathbb{Z}^2, \quad (5.6)$$

where $\{M_{u, v}(da), (u, v) \in \mathbb{Z}^2\}$ are i.i.d. copies of an α -stable random measure M on $\mathbf{A} := \{a_{t, s} \in [0, 1), \sum_{|t|=|s|=1} a_{t, s} < 1\} \subset \mathbb{R}^4$ with control measure equal to the (mixing) distribution Φ of the random vector $a = (a(t, s), |t| = |s| = 1)$ taking values in \mathbf{A} . The random field $\{\mathfrak{X}(t, s)\}$ in (5.6) is α -stable and a particular case of mixed stable moving-average fields introduced in Surgailis *et al.* (1993).

It is not surprising that dependence properties of the random field $\{\mathfrak{X}(t, s)\}$ in (5.6) strongly depend on the concentration of Φ near the “unit root boundary” $A = \sum_{|t|=|s|=1} a_{t, s} = 1$ but also on the form of the autoregressive operator in (5.1). Puplinskaitė and Surgailis (2013) assume that the ‘angular coefficients’ $0 \leq p(t, s) := a(t, s)/A$, $\sum_{|t|=|s|=1} p(t, s) = 1$ are nonrandom and discuss the following three equations

$$X(t, s) = \frac{A}{2} (X(t-1, s) + X(t, s-1)) + \zeta(t, s), \quad (5.7)$$

$$X(t, s) = \frac{A}{3} (X(t-1, s) + X(t, s+1) + X(t, s-1)) + \zeta(t, s), \quad (5.8)$$

$$X(t, s) = \frac{A}{4} (X(t-1, s) + X(t+1, s) + X(t, s+1) + X(t, s-1)) + \zeta(t, s) \quad (5.9)$$

termed the 2N, 3N and 4N models, respectively (N standing for "Neighbour"), with a random 'radial coefficient' $A \in [0, 1]$ having a regularly varying probability density ϕ at $a = 1$:

$$\phi(a) \sim \phi_1(1-a)^\beta, \quad a \nearrow 1, \quad \exists \phi_1 > 0, \quad 0 < \beta < \alpha - 1, \quad 1 < \alpha \leq 2. \quad (5.10)$$

Stationary solution of the above equations in all three cases is given by (5.2), the Green function being written as

$$g(t, s, a) = \sum_{k=0}^{\infty} a^k p_k(t, s), \quad (t, s) \in \mathbb{Z}^2, \quad a \in [0, 1], \quad (5.11)$$

where $p_k(t, s) = P(W_k = (t, s) | W_0 = (0, 0))$ is the k -step probability of the nearest-neighbor random walk $\{W_k, k = 0, 1, \dots\}$ on the lattice \mathbb{Z}^2 with one-step transition probabilities given by $p(t, s) = 1/2(t + s = 1, t \geq 0, s \geq 0)$ (2N), $p(t, s) = 1/3(t + |s| = 1, t \geq 0)$ (3N) and $p(t, s) = 1/4(|t| + |s| = 1)$ (4N), respectively.

Studying long memory properties of mixed moving average random fields $\{\mathfrak{X}_i(t, s)\}, i = 2, 3, 4$ in (5.6) corresponding to model equations (5.7)–(5.9) requires the control of the Green functions in (5.11) as $|t| + |s| \rightarrow \infty$ and $a \rightarrow 1$ *simultaneously* since for any $a_0 < 1$ fixed the $g(t, s, a)$'s, $a < a_0$ decay exponentially fast with $|t| + |s| \rightarrow \infty$. Moreover, the 2N and 3N models exhibit strong anisotropy, characterized by a markedly different scaling behavior from the 4N model. In the Gaussian case $\alpha = 2$ the random fields (5.6) are completely determined by their spectral density $f(x, y)$ in (5.4) and the scaling properties of (5.6) essentially reduce to the low frequency asymptotics of $f(x, y)$ as $(x, y) \rightarrow (0, 0)$. For the lattice isotropic 4N model and $\alpha = 2$ the asymptotics of $f(x, y) \equiv f_4(x, y)$ under (5.10) was studied by Lavancier (2011) (see also Azomahou (2009)):

$$f_4(x, y) = \text{const} \int_0^1 \frac{1}{|1 - (a/4) \sum_{|t|=|s|=1} e^{i(xt+ys)}|^2} \phi(a) da \sim \frac{\text{const}}{(x^2 + y^2)^{1-\beta}}, \quad (x, y) \rightarrow (0, 0).$$

This shows that the large-scale limit of the Gaussian field $\{\mathfrak{X}_4(t, s)\}$ is a fully isotropic self-similar generalized random field on \mathbb{R}^2 . However, for the 2N model the behavior of the spectral density is different

$$f_2(x, y) = \text{const} \int_0^1 \frac{1}{|1 - (a/2)(e^{ixt} + e^{iys})|^2} \phi(a) da \sim \frac{\text{const}}{|x + y|^{1-\beta}} K\left(\frac{(x - y)^2}{|x + y|}\right), \quad (x, y) \rightarrow (0, 0),$$

where $K(u), u \geq 0$ is a bounded continuous function on $[0, \infty)$ with $K(u) \sim \text{const}/u^{1-\beta}, u \rightarrow \infty$ (Lavancier *et al.* (2013)). The above relations can be rewritten as

$$\lim_{\lambda \rightarrow 0} \lambda f_4(\lambda^{1/H} x, \lambda^{1/H} y) = h_4(x, y) := \frac{\text{const}}{(x^2 + y^2)^{1-\beta}}, \quad H = 2(1 - \beta)$$

and

$$\lim_{\lambda \rightarrow 0} \lambda f_2(\lambda^{1/H_1} b_1 x + \lambda^{1/H_2} b_2 y, \lambda^{1/H_1} b_3 x + \lambda^{1/H_2} b_4 y) = h_2(x, y) := \frac{\text{const}}{|x|^{1-\beta}} K\left(\frac{2y^2}{|x|}\right), \quad H_1 = 1 - \beta, \quad H_2 = 2(1 - \beta),$$

where $B = \begin{pmatrix} b_1 & b_2 \\ b_3 & b_4 \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$ is non-degenerated 2×2 -matrix. Note that the limit function h_2 is non-degenerated, in the sense that it depends on both coordinates x and y , and that the scaling limit for f_2 involves two different scaling exponents $H_1 \neq H_2$, as well as a linear transformation of \mathbb{R}^2 with non-degenerated matrix B . Lavancier *et al.* (2013) argue that the above behavior of f_2 is characteristic to anisotropic long memory, in contrast to the isotropic long memory behavior of f_4 .

Puplinskaitė and Surgailis (2013) define *anisotropic distributional long memory* through scaling behavior, or partial sums limits

$$n^{-H_1} \sum_{(t,s) \in \tilde{K}} \mathfrak{X}(t,s) \xrightarrow{\text{fdd}} V(x,y), \quad (x,y) \in \mathbb{R}_+^2 = \{(x,y) \in \mathbb{R}^2 : x,y > 0\} \quad (5.12)$$

on *incommensurate* rectangles $\tilde{K} = K_{[nx, n^{H_1/H_2}y]} := \{(t,s) \in \mathbb{Z}^2 : 1 \leq t \leq nx, 1 \leq s \leq n^{H_1/H_2}y\}$ with sides growing at different rates $O(n)$ and $O(n^{H_1/H_2})$, $H_1 \neq H_2$.

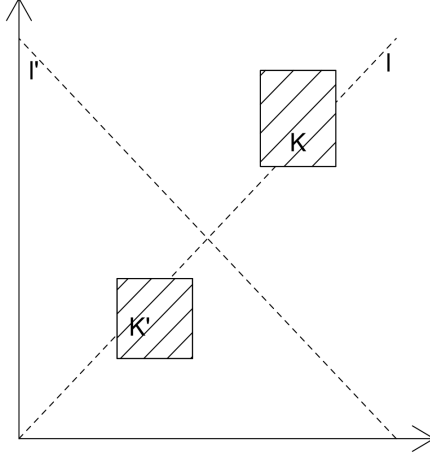


Figure 1: Location of rectangles K and K' .

The limit random field $\{V(x,y)\}$ in (5.12) is assumed to have *dependent increments*, in the following sense. Given a continuous-time random field $V = \{V(x,y)\}$, the *increment of V on rectangle $K = K_{(u,v),(x,y)} = \{(z,w) \in \mathbb{R}^2 : u < z \leq x, u < w \leq y\}$* (with sides parallel to the coordinate axes) is defined as the (double) difference:

$$V(K) := V(x,y) - V(u,y) - V(x,v) + V(u,v).$$

If $\ell = \{(x,y) \in \mathbb{R}^2 : ax + by = c\} \subset \mathbb{R}^2$ is a line, we say that *rectangles K and K' are separated by ℓ* if they lie on different sides of ℓ . Given a line $\ell = \{(x,y) \in \mathbb{R}^2 : ax + by = 0\}$, we say that *V has independent increments in direction ℓ* if for any orthogonal line $\ell' = \{(x,y) \in \mathbb{R}^2 : a'x + b'y = c'\}$, $\ell' \perp \ell$ and any two rectangles K, K' separated by ℓ' , increments $V(K)$ and $V(K')$ are independent r.v.'s; else *V has dependent increments in direction ℓ* (see Fig. 1). Finally, we say that a *random field V has dependent increments* if V has dependent increments in each direction. The above definition of anisotropic distributional long memory is contrasted in Puplinskaitė and Surgailis (2013) with that of *isotropic distributional long memory*, with the only difference that in the latter case, (5.12) is supposed to hold with $H_1 = H_2$, in other words, the rectangles $K_{[nx, ny]}$ grow proportionally in each direction. In both cases, the limit random field satisfies the scaling relation

$$\{\lambda V(x,y)\} \stackrel{\text{fdd}}{=} \{V(\lambda^{1/H_1}x, \lambda^{1/H_2}y)\}, \quad \forall \lambda > 0. \quad (5.13)$$

(5.13) is a particular case of *operator scaling random field* (osrf) introduced in Biermé *et al.* (2007); for $H_1 = H_2$ (5.13) agrees with *self-similar random field* (ssrf). Puplinskaitė and Surgailis (2013) show that the aggregated $2N$ and $3N$ α -stable random fields $\{\mathfrak{X}_2(t,s)\}$ and $\{\mathfrak{X}_3(t,s)\}$ satisfy the above definition of

anisotropic long memory with $H_1 = \frac{(1/2)+\alpha-\beta}{\alpha}$, $H_2 = 2H_1$ for any $1 < \alpha \leq 2$, $0 < \beta < \alpha - 1$, and identify the limiting osrf's V_2 and V_3 as stochastic integrals with respect to an α -stable random measure with integrands involving the limiting Green functions h_2 and h_3 . On the other hand, the 4N field $\{\mathfrak{X}_4(t, s)\}$ is proved to satisfy the isotropic long memory property with $H_1 = H_2 = 2(\alpha - \beta)/\alpha$ and a limiting field V_4 involving the limit Green function $h_4(t, s, z) = \lim_{\lambda \rightarrow \infty} g_4([\lambda t], [\lambda s], 1 - z/\lambda^2)$. The Green functions $h_i, i = 2, 3, 4$ have a classical form of potentials of one-dimensional heat equation and the Helmholtz equation in \mathbb{R}^2 (see Puplinskaitė and Surgailis (2013)).

6 Disaggregation

Aggregation by itself inflicts a considerable loss of information about the evolution of individual “micro-agents”, the latter being largely determined by the mixing distribution. The *disaggregation problem* is to recover the “lost information”, or the mixing distribution from the spectral density or other characteristics of the aggregated series. As a first step in this direction we need to identify the class of spectral densities which arise from aggregation of short memory (SM) processes with random parameters. Dacunha-Castelle and Fermin (2006), Dacunha-Castelle and Oppenheim (2001) obtained an analytic characterization of the class of spectral densities which can be written as mixtures of infinitely differentiable SM spectral densities. Further results in this direction were obtained in Celov *et al.* (2007). For example, the well-known FARIMA(0, d , 0) spectral density $f(y) = (2\pi)^{-1} |2 \sin(|y|/2)|^{-2d}$ can be written as the mixture $f(y) = \int_0^1 \phi(x) |1 - x e^{iy}|^{-2} \phi(x) dx$ of the AR(1) processes corresponding to the mixing density

$$\phi(x) = C(d) x^{d-1} (1-x)^{1-2d} (1+x).$$

See Celov *et al.* (2007) for this and some other examples.

The above disaggregation problem naturally leads to statistical problems such as estimation of the mixing distribution from observed data. The observations may come either from the (limit) aggregated process, or from observed individual processes (sometimes also called panel data). Below we review several methods of estimation of ϕ in the autoregressive aggregation scheme described in Sec. 1. Note that (1.6) means that the covariances of the limit aggregated process are nothing else but the moments of the finite measure having density $(1-x^2)^{-1} \phi(x)$. In fact, being supported by $[0, 1]$, this measure is uniquely determined by its moments, so that the statistical problem of recovering the mixture density from observations $\mathfrak{X}(1), \dots, \mathfrak{X}(n)$ with finite variance is relevant. The last problem is much harder and is still open for the infinite variance aggregated process $\{\mathfrak{X}(t)\}$ defined in (2.4) (Sec. 2).

Estimation from panel data. Consider a panel of N independent AR(1) processes, each of length n . Robinson (1978) and Beran *et al.* (2010) give estimates of ϕ under the assumption that ϕ belongs to some parametric family. The main example is the family of Beta-type distribution of the form

$$\phi_{p,q}(x) = \frac{2}{B(p,q)} x^{2p-1} (1-x^2)^{q-1}, \quad x \in [0, 1], \quad p > 1, q > 1. \quad (6.1)$$

In this context the estimation of ϕ is reduced to the estimation of parameter (p, q) . We outline briefly the two approaches.

Robinson (1978) suggested to use the classical method of moments using the following relation between the moments of ϕ and the auto-covariances $\gamma(k) = EX_j(0)X_j(k)$ (1.3) of individual AR(1) processes,

$$\mu_k = \int_0^1 x^k \phi(x) dx = \frac{\gamma(k) - \gamma(k+2)}{\gamma(0) - \gamma(2)}. \quad (6.2)$$

From panel data, the $\gamma(k)$'s can be estimated as

$$\frac{1}{(n-k+1)N} \sum_{t=1}^{n-k} \sum_{j=1}^N X_j(t) X_j(t+k).$$

Robinson (1978) proved the asymptotic normality of the corresponding estimates of moments $\mu_k, 0 \leq k \leq n$ when N goes to infinity and n is fixed.

Beran *et al.* (2010) proposed an alternative method based on the maximum likelihood estimate (MLE) calculated from estimated observations. The unobserved coefficients a_i of autoregressive processes $\{X_i(t)\}, i = 1, \dots, N$, are estimated by a truncated version of lag-one autocorrelation

$$\hat{a}_{i,n,h} = \min\{\max\{\hat{a}_{i,n}, h\}, 1-h\} \quad \text{with} \quad \hat{a}_{i,n} = \frac{\sum_{t=1}^n X_i(t) X_i(t-1)}{\sum_{t=1}^n X_i^2(t)}, \quad h > 0.$$

In this way we obtain N "pseudo" observations $\hat{a}_{1,n,h}, \hat{a}_{2,n,h}, \dots, \hat{a}_{N,n,h}$ of r.v. a . Then, the parameters p and q in (6.1) are estimated by maximizing the likelihood, viz. $(\hat{p}, \hat{q}) = \arg \max_{p,q} \prod_{i=1}^N \phi_{p,q}(\hat{a}_{i,n,h})$. Beran *et al.* (2010) proved the convergence in probability of the above MLE estimate and its asymptotic normality with the convergence rate \sqrt{N} under the following conditions on the sample sizes and the truncation parameter h : $n, N \rightarrow \infty, h \rightarrow 0, (\log h)^2/\sqrt{N} \rightarrow 0, \sqrt{N}h^{\min(p,q)} \rightarrow 0$ and $\sqrt{N}h^{-2}n^{-1} \rightarrow 0$.

Estimation from the limit aggregated process. Assume that a sample $\mathfrak{X}(1), \dots, \mathfrak{X}(n)$ of size n is observed from the limited aggregated process. Under a parametric assumption about ϕ , the estimator proposed by Robinson (1978) can be easily adapted to this context, because the limit aggregated process and the individual AR(1) have the same covariances.

Leipus *et al.* (2006), Celov *et al.* (2010) use the relation (6.2) to construct a non-parametric estimate of ϕ under the assumption that

$$\phi(x) = (1-x)^{\beta_1} (1+x)^{\beta_2} \psi(x), \quad \beta_1 > 0, \beta_2 > 0, \quad (6.3)$$

where $\psi(a)$ is continuous on $[-1, 1]$ and does not vanishes at $+1, -1$, implying $E(1-a^2)^{-1} < \infty$. The above-mentioned nonparametric estimator is based on the expansion of the mixing density in the orthonormal basis of Gegenbauer's polynomials in the space $L^2(w^{(\alpha)})$ with weight function $w^{(\alpha)}(x) = (1-x^2)^\alpha$ and $\alpha > -1$. The estimate is defined as

$$\hat{\phi}_n(x) := (1-x^2)^\alpha \frac{1}{\sigma^2} \sum_{k=0}^{K_n} \hat{\zeta}_{n,k} G_k^{(\alpha)}(x), \quad (6.4)$$

where the coefficients $\hat{\zeta}_{n,k}$ are defined as follows

$$\hat{\zeta}_{n,k} := \sum_{j=0}^k g_{k,j}^{(\alpha)} (\hat{\gamma}_n(j) - \hat{\gamma}_n(j+2)), \quad (6.5)$$

with $\hat{\gamma}_n(j)$ the sample covariance of the zero mean aggregated process $\{\mathfrak{X}(t), t = 1, \dots, n\}$.

The choice of (K_n) is crucial to obtain consistent estimate of ϕ . Under the condition $\int \frac{\phi(x)^2}{(1-x^2)^\alpha} dx < \infty$, Leipus *et al.* (2006) showed that if K_n is a nondecreasing sequence which tends to infinity at rate $[\gamma \log(n)]$, $0 < \gamma < (2 \log(1 + \sqrt{2}))^{-1}$ then

$$\int_{-1}^1 \frac{E(\hat{\phi}_n(x) - \phi(x))^2}{(1-x^2)^\alpha} dx \rightarrow 0.$$

Celov *et al.* (2010) proved the asymptotic normality $\frac{\hat{\phi}_n(x) - E\hat{\phi}_n(x)}{\sqrt{\text{Var}(\hat{\phi}_n(x))}} \rightarrow_d N(0, 1)$, for every fixed $x \in (-1, 1)$.

The estimate (6.4) depends on the variance $\sigma^2 = \gamma(0) - \gamma(2)$ which can be replaced by its estimate

$\hat{\sigma}^2 = \hat{\gamma}_n(0) - \hat{\gamma}_n(2)$. Philippe *et al.* (2013) proved that the modified estimate is still consistent in a weaker sense, since

$$\int_{-1}^1 \frac{(\hat{\phi}_n(x) - \phi(x))^2}{(1 - x^2)^\alpha} dx \xrightarrow{P} 0.$$

The estimate in (6.4) has been extended non-gaussian aggregated process in (3.4) with finite variance discussed in Sec. 3 (see Philippe *et al.* (2013)) and to some aggregated random field models (see Leonenko and Taufer (2013)).

It should be noted that the estimate (6.4) is not adapted to the limit aggregated process with common innovations. For such models, Chong (2006) proposed a parametric estimate of ϕ , assuming that ϕ is a polynomial. The last condition, however, excludes the case of long memory processes.

References

- Astrauskas, A. (1983). Limit theorems for sums of linearly generated random variables. *Lithuanian Mathematical Journal*, **23**, 127–134.
- Azomahou, T.T. (2009). Memory properties and aggregation of spatial autoregressive models. *Journal of Statistical Planning and Inference*, **139**, 2581–2579.
- Baillie, R. T. (1996). Long memory processes and fractional integration in econometrics. *Journal of Econometrics*, **73**, 5–59.
- Barndorff-Nielsen, O. E. (2001). Superposition of Ornstein–Uhlenbeck type processes. *Theory of Probability and Its Applications*, **45**, 175–194.
- Beran, Feng-Y. Ghosh S. J. and Kulik, R. (2013). *Long-memory processes: Probabilistic properties and statistical methods*. Springer.
- Beran, J., Schützner, M. and Ghosh, S. (2010). From short to long memory: Aggregation and estimation. *Computational Statistics and Data Analysis*, **54**, 2432–2442.
- Biermé, H., Meerschaert, M. M. and Scheffler, H. P. (2007). Operator scaling stable random fields. *Stochastic Processes and Applications*, **117**, 312–332.
- Celov, D., Leipus, R. and Philippe, A. (2007). Time series aggregation, disaggregation and long memory. *Lithuanian Mathematical Journal*, **47**, 379–393.
- Celov, D., Leipus, R. and Philippe, A. (2010). Asymptotic normality of the mixture density estimator in a disaggregation scheme. *Journal of Nonparametric Statistics*, **22**, 425–442.
- Chong, T. T. (2006). The polynomial aggregated AR(1) model. *Econometrics Journal*, **9**, 98–122.
- Cox, D. R. (1984). Long-range dependence: A review. In *Statistics: An Appraisal* (eds H. A. David and H. T. David), pp. 55–74. Iowa State Univ. Press, Iowa.
- Dacunha-Castelle, D. and Fermin, L. (2006). Disaggregation of long memory processes on C^∞ class. *Electronic Communications in Probability*, **11**, 35–44.
- Dacunha-Castelle, D. and Oppenheim, G. (2001). Mixtures, aggregation and long-memory. Prepublications, Université de Paris-Sud Mathématiques.

- Ding, Z. and Granger, C. W. J. (1996). Modeling volatility persistence of speculative returns: a new approach. *Journal of Econometrics*, **73**, 185–215.
- Dombry, C. and Kaj, I. (2011). The on-off network traffic model under intermediate scaling. *Queueing Systems*, **69**, 29–44.
- Gaigalas, R. (2006). A Poisson bridge between fractional Brownian motion and stable Lévy motion. *Stochastic Processes and Applications*, **116**, 447–462.
- Gaigalas, R. and Kaj, I. (2003). Convergence of scaled renewal processes and a packet arrival model. *Bernoulli*, **9**, 671–703.
- Giraitis, L., Koul, H. L. and Surgailis, D. (2012). *Large Sample Inference for Long Memory Processes*. Imperial College Press, London.
- Giraitis, L., Leipus, R. and Surgailis, D. (2010). Aggregation of random coefficient GLARCH(1,1) process. *Econometric Theory*, **26**, 406–425.
- Gonçalves, E. and Gouriéroux, C. (1988). Aggrégation de processus autoregressifs d'ordre 1. *Annales d'Economie et de Statistique*, **12**, 127–149.
- Granger, C. W. J. (1980). Long memory relationship and the aggregation of dynamic models. *Journal of Econometrics*, **14**, 227–238.
- Heyde, C. C. and Yang, Y. (1997). On defining long-range dependence. *Journal of Applied Probability*, **34**, 939–944.
- Jirak, M. (2011). Asymptotic behavior of weakly dependent aggregated processes. *Periodica Mathematica Hungarica*, **62**, 39–60.
- Kazakevičius, V., Leipus, R. and Viano, M.-C. (2004). Stability of random coefficient ARCH models and aggregation schemes. *Journal of Econometrics*, **120**, 139–158.
- Lamperti, J. (1962). Semi-stable stochastic processes. *Transactions of American Mathematical Society*, **104**, 62–78.
- Lavancier, F. (2007). Invariance principles for non-isotropic long memory random fields. *Statistical Inference for Stochastic Processes*, **10**, 255–282.
- Lavancier, F. (2011). Aggregation of isotropic random fields. *Journal of Statistical Planning and Inference*, **141**, 3862–3866.
- Lavancier, F., Leipus, R. and Surgailis, D. (2013). Aggregation of anisotropic random-coefficient autoregressive random field. Preprint.
- Leipus, R., Oppenheim, G., Philippe, A. and Viano, M.-C. (2006). Orthogonal series density estimation in a disaggregation scheme. *Journal of Statistical Planning and Inference*, **136**, 2547–2571.
- Leipus, R., Paulauskas, V. and Surgailis, D. (2005). Renewal regime switching and stable limit laws. *Journal of Econometrics*, **129**, 299–327.
- Leipus, R. and Surgailis, D. (2003). Random coefficient autoregression, regime switching and long memory. *Advances in Applied Probability*, **35**, 1–18.

- Leipus, R. and Viano, M.-C. (2002). Aggregation in ARCH models. *Lithuanian Mathematical Journal*, **42**, 54–70.
- Leonenko, N. and Taufer, E. (2013). Disaggregation of spatial autoregressive processes. *Spatial Statistics*, **3**, 1–20.
- Lobato, I. N. and Savin, N. E. (1998). Real and spurious long-memory properties of stock-market data (with comments). *Journal of Business & Economic Statistics*, **16**, 261–283.
- Mikosch, T. (2003). Modeling dependence and tails for financial time series. In *Extreme Values in Finance, Telecommunications and the Environment* (eds B. Finkenstädt and H. Rootzén), pp. 185–286. Chapman & Hall.
- Mikosch, T., Resnick, S., Rootzén, H. and Stegeman, A. (2002). Is network traffic approximated by stable Lévy motion or fractional Brownian motion? *Annals of Applied Probability*, **12**, 23–68.
- Mikosch, T. and Samorodnitsky, G. (2000). Ruin probability with claims modeled by a stationary ergodic stable process. *Annals of Probability*, **28**, 1814–1851.
- Oppenheim, G. and Viano, M.-C. (2004). Aggregation of random parameters Ornstein-Uhlenbeck or AR processes: some convergence results. *Journal of Time Series Analysis*, **25**, 335–350.
- Paulauskas, V. (2013). On α -covariance, long, short and negative memories for sequences of random variables with infinite variance. Preprint.
- Perilioğlu, K. and Puplinskaitė, D. (2013). Asymptotics of the ruin probability with claims modeled by α -stable aggregated AR(1) process. *Turkish Journal of Mathematics*, **37**, 129–138.
- Philippe, A., Puplinskaitė, D. and Surgailis, D. (2013). Contemporaneous aggregation of a triangular of random coefficient AR(1) processes. Preprint.
- Pilipauskaitė, V. and Surgailis, D. (2013). Large-scale temporal and contemporaneous aggregation of random-coefficient ar(1) processes. Preprint.
- Pipiras, V., Taqqu, M. S. and Levy, L. B. (2004). Slow, fast, and arbitrary growth conditions for renewal-reward processes when both the renewals and the rewards are heavy-tailed. *Bernoulli*, **10**, 121–163.
- Puplinskaitė, D. and Surgailis, D. (2009). Aggregation of random coefficient AR(1) process with infinite variance and common innovations. *Lithuanian Mathematical Journal*, **49**, 446–463.
- Puplinskaitė, D. and Surgailis, D. (2010). Aggregation of random coefficient AR(1) process with infinite variance and idiosyncratic innovations. *Advances in Applied Probability*, **42**, 509–527.
- Puplinskaitė, D. and Surgailis, D. (2013). Aggregation of autoregressive random fields and anisotropic long memory. Preprint.
- Rajput, B. S. and Rosinski, J. (1989). Spectral representations of infinitely divisible processes. *Probability Theory and Related Fields*, **82**, 451–487.
- Robinson, P. (1978). Statistical inference for a random coefficient autoregressive model. *Scand. J. Statist.*, **5**, 163–168.

- Samorodnitsky, G. and Taqqu, M. S. (1994). *Stable Non-Gaussian Random Processes*. Chapman & Hall, New York.
- Sato, K.-I. (1999). *Lévy Processes and Infinitely Divisible Distributions*. Cambridge University Press, Cambridge.
- Surgailis, D., Rosinski, J., Mandrekar, V. and Cambanis, S. (1993). Stable mixed moving averages. *Probability Theory and Related Fields*, **97**, 543–558.
- Teyssière, G. and Kirman, A. P. (2007) (eds). *Long Memory in Economics*. Springer, Berlin, Heidelberg.
- Ward, L. M. and Greenwood, P. E. (2007). $1/f$ noise. *Scholarpedia*, **2**(12), 1537.
- Willinger, W., Taqqu, M. S., Sherman, R. and Wilson, D. V. (1997). Self-similarity through high-variability: statistical analysis of Ethernet LAN traffic at the source level. *IEEE/ACM Transactions on Networking*, **5**, 71–86.
- Zaffaroni, P. (2004). Contemporaneous aggregation of linear dynamic models in large economies. *Journal of Econometrics*, **120**, 75–102.
- Zaffaroni, P. (2007a). Aggregation and memory of models of changing volatility. *Journal of Econometrics*, **136**, 237–249.
- Zaffaroni, P. (2007b). Contemporaneous aggregation of GARCH processes. *Journal of Time Series Analysis*, **28**, 521–544.